Empirical Modeling of an *In Vitro* Activity of Polychlorinated Biphenyl Congeners and Mixtures

David J. Svendsgaard,[1] Thomas R. Ward,[2] Hugh A. Tilson,[2] and Prasada Rao S. Kodavanti[2]

[1]Statistical Consultant, Calabash, NC 28467 USA; [2]Neurotoxicology Division, National Health and Environmental Effects Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711 USA

The goal of this research is to predict an *in vitro* activity of polychlorinated biphenyl (PCB) congeners and their mixtures and to describe the relationship between this activity and chemical structure. The test system used multiple PCB concentrations on each cell culture plate in a repeated measures design, which improved precision for comparing between concentration levels. A weighted regression that accounted for this experimental design feature was used in fitting a nonlinear dose–response exponential model to the PCB concentration–activity data from an *in vitro* test system in which ^{3}H-phorbol ester binding was measured in cerebellar granule cells exposed to different PCB congeners to test for their effects on protein kinase C translocation. The model allowed for the minimum level to be less than control, a common slope, and the estimation of the log of the concentration that produces an activity 50% above the control activity (E50) for 36 congeners and 3 commercial mixtures. Next, a weighted logistic regression using a second order response model in the variables Cl_{ortho}, Cl_{para}, and Cl_{meta} was used to relate the estimated log E50s to indicators of chemical structure. This model was preferred over models that might seem more mechanistically based because in internal validation, it attained a smaller PRESS statistic (the sum of squares between all observed and predicted observations) than other models. Evidently, this second order model makes more efficient use of parameters than other models considered. Plots of the predictions of the logistic second order response model versus log K_{ow} confirm the usual pattern that congeners with intermediate levels of log K_{ow} are the more active. The data of three commercial mixtures were included in this regression by assuming a common combination index (ratio of observed E50 to predicted E50, assuming dose addition). The logistic model suggests that congeners with one, two, or three chlorine substitutions at the *ortho* position are more active than other congeners. Also, congeners with log K_{ow} between 5.2 and 6.6 are generally more active. The estimated combination index indicated that the joint action of PCB congeners in the three commercial mixtures was less than dose additive. The error sum of squares was significantly large, which may indicate a lack of fit of the logistic model. Empirical Bayes estimates (EBE) are weighted averages of model predictions and observations of E50s and can be better estimates than the fitted model when there is a lack of fit. The PRESS statistic for the EBE indicated larger prediction error than for the logistic model, but the EBE provided better estimates of commercial mixture E50s based on dose addition. This may indicate that the logistic model is not incorporating all the information in the single congener data needed to predict mixtures. *Key words*: empirical model, *in vitro* activity, neurotoxicity, polychlorinated biphenyls, repeated measures design.

Environ Health Perspect 105:1106–1115 (1997). *http://ehis.niehs.nih.gov*

Polychlorinated biphenyls (PCBs) are industrial compounds detected in air, water, sediments, fish, wildlife, and humans (*1,2*). PCBs are prepared by the chlorination of biphenyls, and the result is a mixture of possibly as many as 209 congeners. Although the manufacture of PCBs has been banned in the United States, these compounds remain a serious environmental pollutant due to ongoing release from hazardous waste sites, the accidental breakdown of electric transformers, and the high resistance to degradation.

Risk assessment of PCBs currently involves usage of toxic equivalency factors (TEFs), which are predicted on the assumption that this class of chemicals elicits their toxic responses through a common receptor-mediated mechanism. Structure–activity relationship studies, for example, have found that PCB-induced body weight loss, thymic atrophy, immunotoxicity, endocrine and reproductive toxicity, and carcinogenicity are associated with high affinity for the aryl hydrocarbon (Ah) receptor (*3*). Recent studies, however, have shown environmentally relevant *ortho*-substituted PCB congeners with weak or no Ah-receptor activity have effects on brain dopamine concentrations *in vivo* (*4*) and *in vitro* in PC12 cells (*5*). Kodavanti et al. (*6*) have also found that *ortho*-substituted PCB congeners have significant effects on calcium homeostasis mechanisms *in vitro*, while non-*ortho* PCBs, having a more coplanar structural configuration, are relatively inactive *in vitro*.

This paper builds on the previous work of Kodavanti et al. (*6*) by developing an empirical model for predicting the *in vitro* activity of PCB congeners. This information may be useful in developing a risk assessment strategy for nondioxinlike PCBs and their mixtures. The activity was measured in an *in vitro* test system in which different PCB congeners and PCB mixtures were tested for their effects on protein kinase C (PKC) translocation by measuring ^{3}H-phorbol ester binding in cerebellar granule cells (*6,7*). The prediction of the activity of PCB mixtures is based upon knowledge of the chemical composition and the *in vitro* activity of the components of the mixture.

We considered using the logistic dose–response model to determine the effective concentration that produces an activity 50% above the control activity (E50s) for 36 tested congeners and the 3 commercial mixtures. The logistic model involves the estimation of a minimum and a maximum activity. However, estimates of the maximum were quite variable because not all congeners attained maximum activity. The exponential model is similar to the logistic model but has no maximum value so this model was used instead. Logs of the mean activity relative to control activity were fit to an exponential model using the data from 36 congeners and the 3 commercial mixtures; the model had a common slope parameter, a common minimum value, and a separate log E50 for each of the congeners and the commercial mixtures. There is a correlation between the observed activity values from the same congener due to variability between culture plates. The use of weighted regression based on variance components estimated from the repeated measures design permitted data from all congeners to be estimated simultaneously while addressing this

Address correspondence to D.J. Svendsgaard, 8874 Lansdowne Drive NW, Calabash, NC 28467 USA.

Some of this work was done while D.S. was a member of the Health Effects Research Laboratory. Chris Waller provided valuable insight to this project. We thank Mats Tysklind of Umea University, Sweden, and Subash Basak of National Resources Research Institute, Duluth, Minnesota, for their excellent comments on the earlier version of this manuscript.

The research described in this article has been reviewed by the National Health and Environmental Effects Research Laboratory, U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the agency nor does the mention of trade names or commercial products constitute endorsement or recommendation for use.

Received 27 December 1996; accepted 13 August 1997.

correlation. The weights formed a block diagonal matrix. The minimum value of the dose–response curve was allowed to be negative (less than control) in view of the recent interest in U-shaped dose–response curves (*8*).

Next, a structure–activity relationship needed to be developed so that the E50s of the untested congeners could be estimated. After considering various models involving log K_{ow} and indicators of chemical structure, we decided to approximate this relationship by a logistic second order response surface model in the variables Cl_{ortho}, Cl_{para}, and Cl_{meta}. This model was fit to the logs of the E50s using weighted nonlinear regression. The logs of the E50s of the three commercial mixtures were simultaneously fitted in this regression through the use of the dose-addition equation. For these mixtures, a single combination index was also estimated. The combination index is the ratio of the E50 of the mixture to the E50 predicted by the dose-addition assumption. The reason for simultaneously estimating the structural activity relationship and the mixture E50s was to constrain the estimated structural relationship to one capable of predicting the activity of mixtures based upon dose additivity.

It is ideal to develop a model with one set of data and validate the model with a second set of data; this validation process is called external validation. In this case, all the data were needed to develop the model due to the small sample size; therefore, the model was internally validated. Internal validation involves excluding one data point at a time from the model fitting process and comparing the resultant prediction of the excluded point with the observation of that point. After each observation has been excluded and predicted, the sum of squares of the prediction differences is called the PRESS statistic. The lower the PRESS statistic the better the model.

The PRESS statistic was calculated for five models, and most of them were more biologically based than the second order model. Hansch, the father of quantitative structure–activity relationships (QSAR), derived a quadratic polynomial to relate biological activity to log P (*9*). Therefore, one model we considered was a quadratic model in log K_{ow}. The quadratic coefficient was fixed and the linear coefficients were allowed to vary with Cl_{ortho}. The intercepts varied with the main effects of Cl_{ortho} and Cl_{para}. There were biological reasons for believing that Cl_{ortho} is an important variable, but the reason for considering Cl_{para} was more empirically based. A main effects model was also fit using dummy variables that considered Cl_{ortho}, Cl_{para}, and Cl_{meta} to be factors. The second order logistic model attained a smaller PRESS value than other models.

In choosing the model we were guided by the PRESS statistic and how well the E50s of the mixtures were predicted. The data contain more information about the activity and chemical structure relationship than we have modeled. This state of affairs is due to our search for the most detailed model that could be validated with this sample size. Also, the PRESS statistic seems to favor models that use continuous independent variables rather than dummy variables such as would be used to describe a main effect model in the factors CL_{ortho}, Cl_{para}, and Cl_{meta}.

These efforts are preliminary in some ways. The use of physicochemical variables as independent variables may further explain these activity data. However, considering the prevalence of mixtures in the environment, our focus is on the prediction of the activity of mixtures. This work is a step in that direction, and it gives a framework for more detailed structural-activity research to build on.

Our avoidance of physicochemical variables at this stage of the research is based upon three reasons. First, the selection of congeners to test was based mainly on their presence in environmental mixtures; therefore, the sample of congeners is unlikely to be ideal for the determination of a mechanistically based model. Second, there is measurement error in these variables. Sabljić (*10*) has described the measurement error in log K_{ow} and has concluded that the use of the *n*-octanol/water partition should be avoided in environmental research. Approaches using physicochemical variables as independent variables in regression routines need to address this error. The practice of combining several physicochemical variables into a summary variable using principle component analysis does address the measurement error problem. Third, most biologists do not think in terms of physicochemical variables and describe their results in terms of indices of structure, such as we use here. For example, Kodavanti et al. (*11*) has found that noncoplanar congeners are usually more active. Noncoplanarity is related to the configuration of the *ortho* substitutions. Our descriptive approach sheds light on the importance of such theories in explaining the variation of the E50 estimates.

The success of our approach depends on an assumption of smoothness of the E50 response surface. For example, the logistic second order model can estimate at most one peak in a response surface that might not be very smooth. Our approach should be judged on its clarification of important aspects of the E50 response surface, on its ability to predict mixtures, and on its ability to indicate how future experiments can be designed to improve mixture prediction. We also present some results based upon empirical Bayesian analysis. This approach can fit response surfaces that are less smooth than the logistic second order model can fit.

Materials and Methods

The* in vitro *test system. The *in vitro* test system of Kodavanti et al. (*6,7*) measures increases in [^{3}H]-phorbol ester (PDBu) binding, which suggests increased activation/translocation of PKC from cytosol to the membrane. Translocation of PKC is dependent on the concentrations of intracellular free Ca^{2+} and/or diacylglycerol (*12,13*). PKC has been reported to play a key role in a number of physiological and toxicological phenomena (*14,15*). Cerebellar granule cells grown on 12-well culture plates were tested after 7 days in culture for [^{3}H]-PDBu binding. Each replicate consisted of a control and usually six different concentrations placed in the wells of the cell culture plates. Generally, four replicates were used and, other than control, the concentrations of PCBs were 1, 3, 10, 30, 50, and 100 μM. Table 1 shows the details of the experimental designs.

Chemicals and terminology. Figure 1 shows the chemical structure of a PCB congener and explains how the congeners are denoted. The log of the octanol–water partition coefficient is denoted by log K_{ow} and was obtained from Hawker and Connell (*16*). The tested congeners are listed in Table 1. The percentages in the commercial mixtures were the averages across lots of values from Frame et al. (*17*).

The concentration–activity model. The $\log_{10}$ of activity ([^{3}H]-PDBu binding) relative to control and averaged across culture plates was fit to an exponential concentration–activity model. Specifically, the model for the *j*th congener was

$$y = m_0 + \left(k - m_0\right)e^{s\left[\left(\log_{10} C\right) - \log_{10} \mathrm{E50}_j\right]} \quad (1)$$

where y is the average across plates of the $\log_{10}$ of [^{3}H]-PDBu binding relative to the control activity, C is the concentration (μM), and $k = \log_{10} 1.5$. The parameter m_0 is the log of the minimum of the concentration–activity model, and s is a slope parameter. This parameterization allows the standard error of the log E50 to be obtained directly from the nonlinear routine, which is based on the delta method.

This model was fit simultaneously to all the nonzero dose data from the 36 congeners and 3 commercial mixtures using weighted regression. The weights were determined as follows. A repeated measures analysis was done on the log of the activity relative to control for each congener. This resulted in

Table 1. Experimental conditions

IUPAC	Structure	Log K_{ow}	Plates	Low concentration (μM)	High concentration (μM)	Number of doses
4	2,2′	4.65	5	1	100	6
11	3,3′	5.28	4	1	100	6
14	3,5	5.20	4	1	100	6
15	4,4′	5.30	4	1	100	6
19	2,2′,6	5.02	4	1	100	6
21	2,3,4	5.51	4	3	100	6
28	2,4,4′	5.67	6	1	100	6
47	2,2′,4,4′	5.85	4	1	100	6
50	2,2′,4,6	5.63	4	1	100	6
51	2,2′,4,6′	5.63	4	1	100	6
52	2,2′,5,5′	5.84	4	1	100	6
53	2,2′,5,6′	5.62	4	3	100	5
54	2,2′,6,6′	5.21	4	1	100	6
77	3,3′,4,4′	6.36	4	1	100	6
80	3,3′,5,5′	6.48	4	1	100	6
82	2,2′,3,3′,4	6.20	4	10	100	4
85	2,2′,3,4,4′	6.30	4	10	100	4
95	2,2′,3,5′,6	6.13	4	10	100	4
99	2,2′,4,4′,5	6.39	4	10	100	4
100	2,2′,4,4′,6	6.23	4	3	100	5
101	2,2′,4,5,5′	6.23	4	10	100	4
104	2,2′,4,6,6′	5.81	4	1	100	6
105	2,3,3′,4,4′	6.65	5	1	100	6
110	2,3,3′,4′,6	6.48	4	10	100	4
118	2,3′,4,4′,5	6.74	4	1	100	6
126	3,3′,4,4′,5	6.89	4	3	100	4
127	3,3′,4,5,5′	6.95	4	3	100	5
128	2,2′,3,3′,4,4′	6.74	4	1	100	6
133	2,2′,3,3′,5,5′	6.86	3	1	100	6
136	2,2′,3,3′,6,6′	6.22	4	1	100	6
138	2,2′,3,4,4′,5′	6.83	4	10	100	4
153	2,2′,4,4′,5,5′	6.92	4	1	100	6
156	2,3,3′,4,4′,5	7.18	5	1	100	6
169	3,3′,4,4′,5,5′	7.42	4	1	100	6
179	2,2′,3,3′,5,6,6′	6.73	4	10	100	4
180	2,2′,3,4,4′,5,5′	7.36	4	1	100	6
AR1016	Mix	5.49[a]	5	1	100	6
AR1254	Mix	6.51[a]	4	1	100	6
AR1260	Mix	6.98[a]	5	1	100	6

IUPAC, International Union of Pure and Applied Chemistry.
[a]Based on average of congeners present.

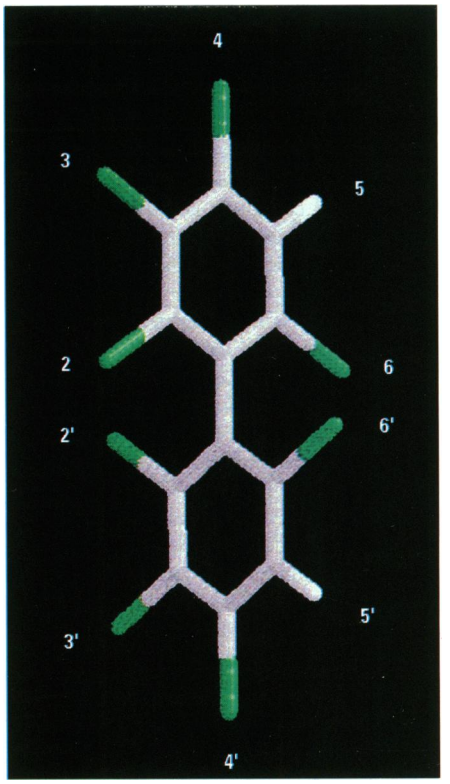

Figure 1. The chemical configuration of polychlorinated biphenyls (PCBs) for 2,2′,3,3′,4,4′,6,6′. Green indicates the substitution of chlorine at the position and white indicates an unsubstituted position. PCB congeners have two rings, which can rotate around the bond connecting them. On each ring there are five positions where chlorine substitution can occur. These positions are numbered as shown. The *ortho* positions are 2,2′,6, and 6′; the *meta* positions are 3,3′,5, and 5′; and the *para* positions are 4 and 4′.

estimates for the variability due to plates, variability due to doses, and residual variability.

For the data from each congener (or mixture), a covariance matrix with plate variability at the off-diagonal elements and plate plus residual variability on the diagonal was formed. The weights for all the 36 congeners and the 3 commercial mixtures formed a block diagonal matrix. See Draper and Smith (*18*) for an explanation of why it is valid to use linear regression with such a weight matrix.

The dose-addition method of predicting the mixture E50. The E50 of a mixture satisfies the following equation when the joint action is dose addition:

$$\frac{1}{\mathrm{E50}_k} = \sum_{j=1}^{209} \frac{\pi_{jk}}{\mathrm{E50}_j}, \qquad (2)$$

where π_{jk} is the fraction of the jth congener in the kth mixture. The combination index (*19*) is the ratio of the actual E50 of the mixture to the E50 predicted in Equation 2. Combination indices that are >1 indicate antagonism and those <1 indicate synergism.

The structure–activity model. Weighted nonlinear regression analysis was used to determine the relationship between the log of the E50 of the concentration–activity model and the chemical structure of the PCB congener. Also, the logs of the E50 for the three commercial mixtures were included in this regression by assuming that these mixtures all had the same combination index. The weights used were the reciprocal of the square of the estimated standard errors of the log E50s. The model used to relate the log E50 of each congener to the structure was the second order logistic model:

$$\log \mathrm{E50} = m + (M-m)/(1 + \exp[\gamma_0 + \gamma_1(Cl_{ortho}-2) + \gamma_2(Cl_{para}-1) + \gamma_3(Cl_{meta}-2) + \gamma_{11}(Cl_{ortho}-2)^2 + \gamma_{22}(Cl_{para}-1)^2 + \gamma_{33}(Cl_{meta}-2)^2 + \gamma_{12}(Cl_{ortho}-2)(Cl_{para}-1) + \gamma_{13}(Cl_{ortho}-2)(Cl_{meta}-2) + \gamma_{23}(Cl_{para}-1)(Cl_{meta}-2)] \qquad (3)$$

The logistic model becomes flat as the value of the expression within the square brackets of the equation becomes either positively or negatively large. This feature restricts the predicted values to lie in the range from m to M (minimum to maximum), thus improving the stability of the predictions.

The prediction of mixture E50s. Using the combination index and Equation 2, the log of the E50 of the mth mixture can be expressed as

$$\log \mathrm{E50}_m = \log CI_m - \log \sum_{i=1}^{209} \pi_i 10^{-\log \mathrm{E50}_i} \qquad (4)$$

where all logs are to the base 10 and CI_m is the combination index for the mth mixture. In fact, when one views each congener as a

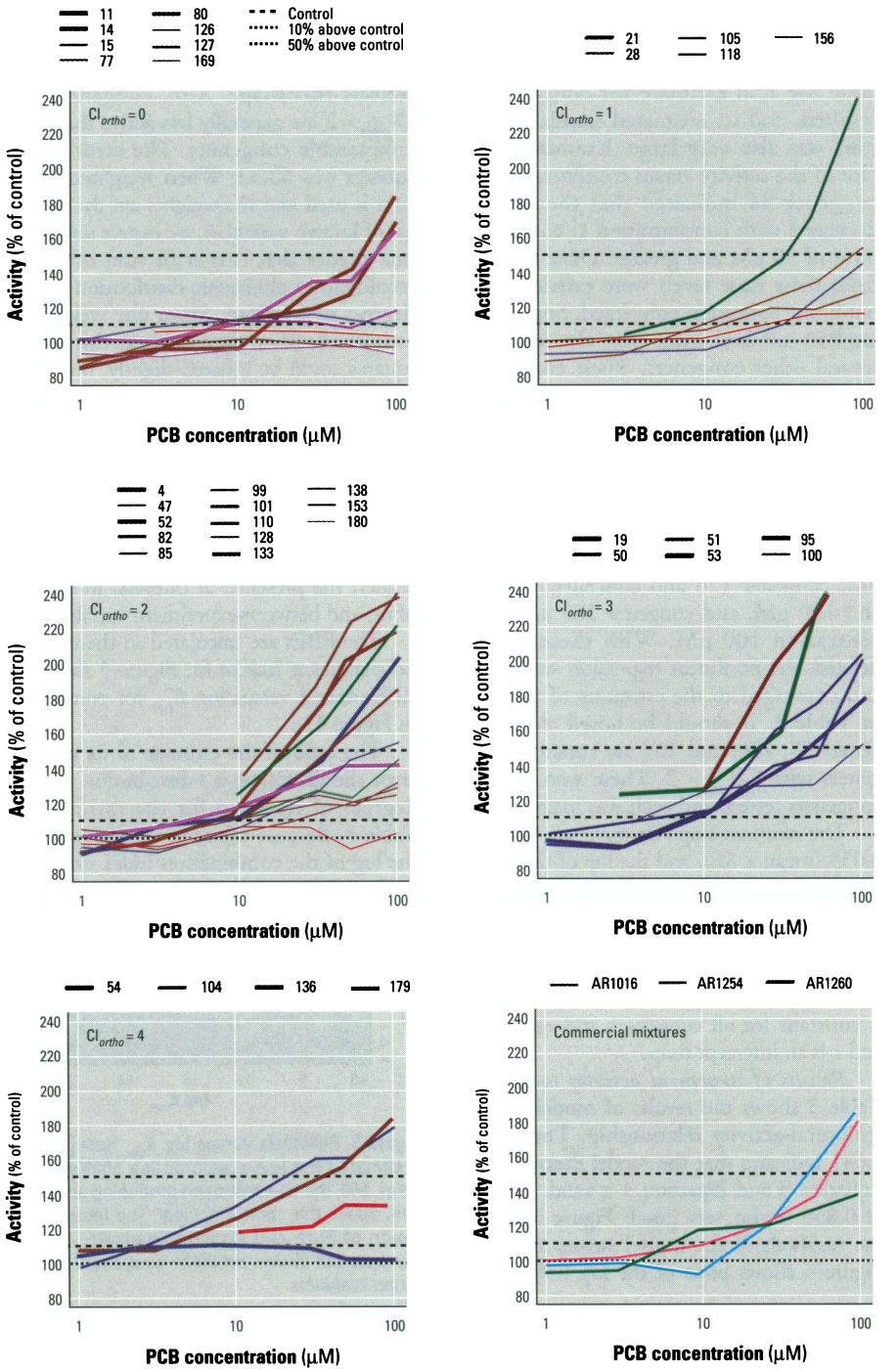

Figure 2. Activity–concentration curves. These curves show the average activity in percent control versus polychlorinated biphenyls (PCB) concentration for each of the congeners and the commercial mixtures. The color of the activity–concentration line represents the level of Cl_{meta}: blue is 0, green is 1, brown is 2, orange is 3, and purple is 4. The wider the activity–concentration line, the less chlorine at the *para* position. IUPAC, International Union of Pure and Applied Chemistry.

mixture having 100% of one congener (π_i = 1) and CI_m = 1, then the log of the E50s of the 36 tested congeners also satisfies Equation 4. The equation fit to the data was Equation 4, with Equation 3 substituted for log $E50_i$ in Equation 4.

Cross-validation. The PRESS statistic is the sum of squares of differences between observed and predicted and is used for internal validation. The predicted are obtained for each observation by leaving out that observation from the regression routine, using the rest of the observations to fit the regression model, and predicting the deleted observation with the resultant fitted value. The Q^2 statistic is 1–PRESS/$PRESS_0$ where $PRESS_0$ is the PRESS statistic for an overall mean model using the same dependent variable. The Q^2 statistic is similar to R^2. When Q^2 is 1, the method has excellent internal validation. Unlike R^2, Q^2 can become negative for poor predictors. The Q^2 statistic can also be computed for sets of data not used in the regression, and that statistic is called the Q^2 for external validation. A weighted Q^2 was also computed by weighting the sum of squares by the same weights used in the regressions.

Empirical Bayesian analysis. Bryk and Raudenbush (*20*) discuss empirical Bayesian analysis. Briefly, when data are analyzed by using a simple weighted linear regression with one continuous independent variable, the error sum of squares has a chi-square distribution, with degrees of freedom equal to the residual degrees of freedom if several conditions are met. These are 1) the elements of the dependent variable are normally distributed with a known variance, 2) the reciprocal of this variance is used as the weight, and 3) the dependent variable is linearly related to the independent variable. When this is the case, the empirical Bayes estimates (EBEs) of the true model can be used; they are points lying between the elements of the dependent variable and the straight line fitted by the regression. The smaller the error sum of squares, the closer the EBE is to the straight line.

We used the EBE in the case of weighted nonlinear regression with multiple independent variables. The weights must be estimated because the standard errors are unknown. Therefore, we expect the error sum of squares to be somewhat more variable than a chi-square distribution.

We tested 36 congeners so we could estimate their E50s directly. The E50 of the 173 untested congeners could be estimated from the fitted second order logistic model. By assuming there were indications of lack of fit of the model, we could use the EBEs in this case. However, EBEs only are known for the 36 tested congeners. To estimate the EBEs for the untested congeners, linear splines were connected to the EBE points to estimate the EBEs for the untested congeners. Those untested congeners with log K_{ow} values less (or greater) than the smallest (largest) tested congener were assigned the same EBE as the smallest (or largest) tested congener. These six groups were the coplanar congeners with Cl_{para} = 0; the coplanar congeners with Cl_{para} = 1 or 2; the congeners with Cl_{ortho} = 1; the congeners with Cl_{ortho} = 2; the congeners with Cl_{ortho} = 3; and the congeners with Cl_{ortho} = 4.

The EBE for the ith observation ($observed_i$) were obtained as follows:

$$EBE_i = \lambda_i observed_i + (1 - \lambda_i) fitted_i, \quad (5)$$

where $\lambda_i = \tau/(\tau + \upsilon_i)$, τ is the unexplained variance between log E50s, υ_i is the estimated variance of the log E50 of the ith congener, and fitted$_i$ is the predicted value for the ith observation. The variance τ was estimated by adding a constant to the reciprocal of the weights. This constant was increased until the error sum of squares was less than the 90th percentile of a chi-square with the residual degrees of freedom. This constant was used as the estimate of τ.

Results

Results of fitting an exponential concentration–activity model to 36 congeners and 3 commercial mixtures. Figure 2 shows the mean activity versus concentration over the region from 1 to 100 µM for the 36 congeners and 3 commercial mixtures. Note that 22 of the congeners have mean activities exceeding 50% of control. About 6 of the congeners barely exceed 10% of control.

Table 2. Estimates of log base 10 E50s from the exponential model

IUPAC	log E50	SE of log E50
4	1.531	0.111
11	1.488	0.085
14	1.848	0.102
15	3.148	0.546
19	1.689	0.140
21	1.354	0.064
28	1.936	0.118
47	1.792	0.113
50	1.504	0.069
51	1.467	0.081
52	1.341	0.065
53	1.220	0.061
54	1.842	0.353
77	2.832	1.190
80	1.819	0.114
82	1.300	0.099
85	1.412	0.089
95	1.194	0.089
99	2.125	0.143
100	1.985	0.080
101	1.659	0.102
104	1.627	0.112
105	1.808	0.116
110	1.241	0.036
118	2.326	0.170
126	2.055	0.554
127	2.468	0.099
128	2.269	0.199
133	2.077	0.172
136	1.645	0.154
138	2.215	0.142
153	2.247	0.142
156	2.602	0.217
169	3.660	0.837
179	2.104	0.092
180	4.381	1.964
AR1016	1.764	0.104
AR1254	1.798	0.081
AR1260	2.018	0.140

Abbreviations: IUPAC, International Union of Pure and Applied Chemistry; E50, the concentration producing an activity 50% above the control activity; SE, standard error.

A difficulty was encountered when fitting the exponential concentration–activity model (Equation 1) to the data. A very large log E50 estimate for congener 54 resulted, and its estimated standard error (SE) was also very large. Examining the plot of the activity versus concentration for congener 54 indicated that the activity decreased with concentration at concentrations of 30 µM and greater. Once the data from these dose levels were excluded, the nonlinear regression converged. Some large log E50 standard errors were also noted for several other congeners. Their concentration–activity relationships were also plotted, and decreasing activity at increasing concentration was also noted for those congeners. These data points were also excluded. Altogether, 9 of the 212 data points were excluded. These were congeners 54 and 77 and concentrations 30, 50, and 100 µM; congener 126 and concentrations 50 and 100 µM, and congener 169 and concentration 100 µM. With these points deleted, the nonlinear regression was stable and converged to the estimates of log E50 in Table 2. It should be noted that congeners 77, 126, and 169 are coplanar congeners with $Cl_{para} = 2$. There were 4 such congeners tested, and each was inactive.

The resultant slope estimate was 0.494 ± 0.033 (mean ± SE), and the log of the minimum $m_0 = -0.0168 \pm 0.0079$. The model for the ith congener is of the form log activity = $-0.017 + s_i d^{0.494}$, where the parameter s_i can vary among the congeners and mixtures. On a log dose scale, the slope (0.494) is constant for all congeners and mixtures, and s_i is an intercept term.

Results of structural activity modeling. Table 3 shows the results of modeling the structural-activity relationship. The Q^2 of 0.641 indicates that the model described by Equation 3 and Equation 4 is valid. The R^2 of 0.866 is also very good. Figure 3 shows the residuals of the model versus log K_{ow}. Figure 4 shows plots of the log E50 versus log K_{ow}. Generally, we expect to see activity decrease and then increase as log K_{ow} increases. The plots show that the model generally reflects this shape. Also, congeners with $Cl_{para} = 2$ are generally less active than other comparable congeners. The error sum of squares was 68.32. When weighted regression is used and the weights are the reciprocal of known variances, we expect under normal theory that this error sum of squares would have a chi-square distribution with 26 degrees of freedom. Since our weights are estimated rather than known, this sum of squares could be inflated slightly, but not to this amount. A simulation study indicated that using estimated weights based on t-distributions with 21 degrees of freedom would increase the 90th percentile about 10%. The 90th percentile of a chi-square with 26 degrees of freedom is 35.6. Factors that tend to increase this sum of squares are nonnormality, the presence of outliers, model lack of fit, and between-experiment variability.

The EBEs are calculated so there will be no significant lack of fit. Figure 5 shows the EBE plotted versus log K_{ow} for comparison to Figure 4.

The ratio of the estimate to its standard error should follow a t-distribution with 26 degrees of freedom for the estimates in Table 3. Because the ratio corresponding to the log of the combination index exceeds 2,

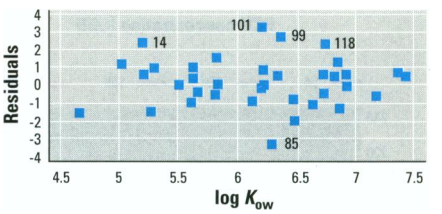

Figure 3. Residuals versus log K_{ow}. Residuals are (estimated log E50-predicted log E50)/standard error. E50, the concentration producing an activity 50% above the control activity. The International Union of Pure and Applied Chemistry (IUPAC) numbers are shown for those congeners with the larger residuals.

Table 3. Results of structural-activity modeling

Parameters	Variables	Estimate	SE	Ratio
m	Minimum	1.142	0.196	5.824
M	Maximum	3.274	1.169	2.801
γ_0	Constant	2.442	0.875	2.790
γ_1	Cl_{ortho}-2	-0.379	0.243	-1.560
γ_2	Cl_{para}-1	-1.473	0.887	-1.661
γ_3	Cl_{meta}-2	-0.774	0.450	-1.720
γ_{11}	$(Cl_{ortho}-2)^2$	-0.404	0.260	-1.555
γ_{22}	$(Cl_{para}-1)^2$	-0.858	0.524	-1.639
γ_{33}	$(Cl_{meta}-2)^2$	-0.544	0.338	-1.612
γ_{12}	$(Cl_{ortho}-2) \times (Cl_{para}-1)$	-0.452	0.332	-1.363
γ_{13}	$(Cl_{ortho}-2) \times (Cl_{meta}-2)$	-0.413	0.276	-1.498
γ_{23}	$(Cl_{para}-1) \times (Cl_{meta}-2)$	-0.684	0.453	-1.511
$\log_{10}$ CI	$\log_{10}$ CI	0.293	0.100	2.931

Abbreviations: SE, standard error; CI, combination index. $R^2 = 0.866$; $Q^2 = 0.641$.

the log of the combination index is significantly different from zero. This indicates that the mixtures are less active than predicted by the logistic model using dose addition.

Analysis of outliers. Two residuals were larger than 3 standard errors of the estimated log E50s. These were congener numbers 85 and 101. A plot of these residuals is shown in Figure 3. No observations were deleted from this regression analysis.

Table 4 shows the model predictions of the log base 10 E50s for all congeners together with the 95% confidence limits. Also shown are the log K_{ow}. Only three congeners were tested whose log K_{ow} exceeded 7. The large confidence limits of congeners with log K_{ow} greater than 7 reflect the fact that the regression model is being extrapolated outside the experimental region to estimate the E50 of such congeners. Also shown in Table 4 are the EBEs for the log E50s. A value of 0.011 was estimated for τ. The Q^2 was 0.159, which is smaller than that obtained for the second order model, and indicates less predictability. To further evaluate the predictability of these estimates, the mixture E50s were calculated from the EBEs. The observed E50 for AR1016, AR1254, and AR1260 was 58.1, 62.9, and

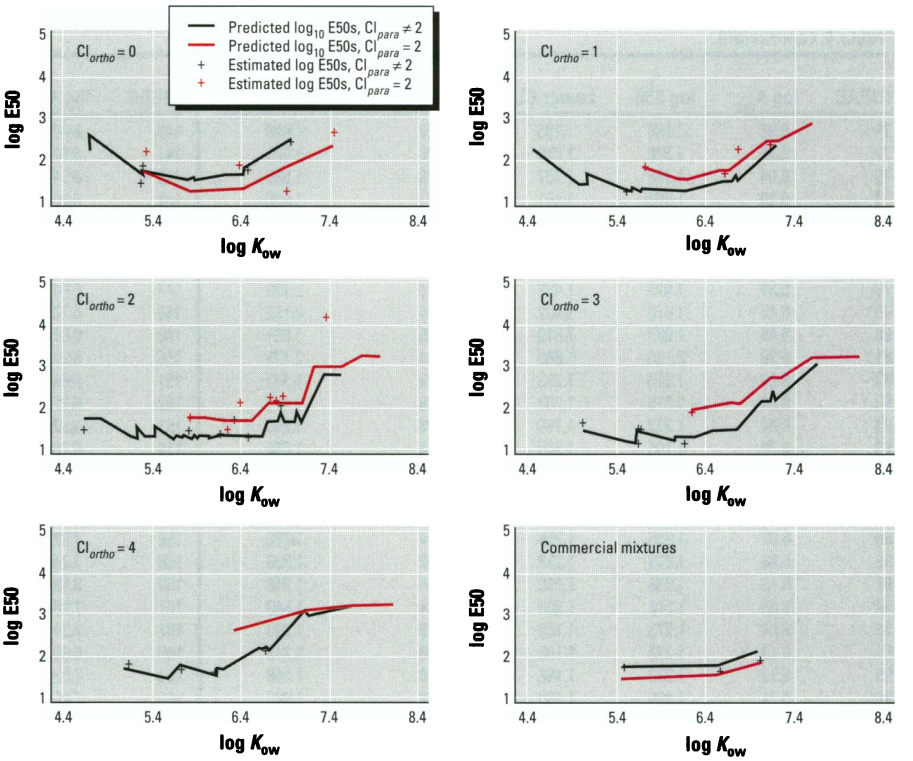

Figure 4. Predicted $\log_{10}$ E50s versus log K_{ow} by CI_{ortho}. E50, the concentration producing an activity 50% above the control activity.

Table 4. Log K_{ow} (lp) and predictions of log E50 and 95% combination index from the logistic second order model

IUPAC	log K_{ow}	log E50	Lower CL	Upper CL	Empirical Bayes estimate
1	4.46	2.338	1.678	2.999	1.370
2	4.69	2.201	1.615	2.787	1.638
3	4.69	2.566	1.848	3.284	2.517
4	4.65	1.691	1.395	1.987	1.603
5	4.97	1.468	1.325	1.611	1.370
6	5.06	1.468	1.325	1.611	1.370
7	5.07	1.732	1.369	2.095	1.370
8	5.07	1.732	1.369	2.095	1.370
9	5.06	1.468	1.325	1.611	1.370
10	4.84	1.691	1.395	1.987	1.558
11	5.28	1.601	1.406	1.796	1.638
12	5.22	1.721	1.367	2.076	2.517
13	5.29	1.721	1.367	2.076	2.517
14	5.28	1.601	1.406	1.796	1.638
15	5.30	2.593	2.006	3.180	2.517
16	5.16	1.289	1.190	1.388	1.483
17	5.25	1.441	1.301	1.582	1.462
18	5.24	1.289	1.190	1.388	1.464
19	5.02	1.514	1.290	1.739	1.580
20	5.57	1.292	1.197	1.387	1.589
21	5.51	1.348	1.253	1.444	1.370
22	5.58	1.348	1.253	1.444	1.625
23	5.57	1.292	1.197	1.387	1.589
24	5.35	1.289	1.190	1.388	1.438
25	5.67	1.348	1.253	1.444	1.954
26	5.66	1.292	1.197	1.387	1.917
27	5.44	1.289	1.190	1.388	1.417
28	5.67	1.972	1.672	2.272	1.954
29	5.60	1.348	1.253	1.444	1.698
30	5.44	1.441	1.301	1.582	1.417
31	5.67	1.348	1.253	1.444	1.954
32	5.44	1.441	1.301	1.582	1.417
33	5.60	1.348	1.253	1.444	1.698
34	5.66	1.292	1.197	1.387	1.917
35	5.82	1.505	1.302	1.708	2.330
36	5.88	1.535	1.358	1.713	1.750
37	5.83	2.079	1.621	2.536	2.326
38	5.76	1.505	1.302	1.708	2.351
39	5.89	1.505	1.302	1.708	2.305
40	5.66	1.238	1.107	1.369	1.365
41	5.69	1.280	1.195	1.365	1.358
42	5.76	1.280	1.195	1.365	1.342
43	5.75	1.238	1.107	1.369	1.344
44	5.75	1.238	1.107	1.369	1.344
45	5.53	1.278	1.165	1.391	1.294
46	5.53	1.278	1.165	1.391	1.294
47	5.85	1.781	1.612	1.949	1.798
48	5.78	1.280	1.195	1.365	1.337
49	5.85	1.280	1.195	1.365	1.798
50	5.63	1.431	1.323	1.538	1.481
51	5.63	1.431	1.323	1.538	1.481
52	5.84	1.238	1.107	1.369	1.323
53	5.62	1.278	1.165	1.391	1.244
54	5.21	1.621	1.237	2.004	1.614
55	6.11	1.317	1.236	1.397	1.931
56	6.11	1.317	1.236	1.397	1.931
57	6.17	1.326	1.216	1.436	1.928
58	6.17	1.326	1.216	1.436	1.928
59	5.95	1.238	1.107	1.369	1.667
60	6.11	1.699	1.491	1.907	1.931
61	6.04	1.317	1.236	1.397	1.935
62	5.89	1.280	1.195	1.365	1.745
63	6.17	1.317	1.236	1.397	1.928
64	5.95	1.280	1.195	1.365	1.667
65	5.86	1.238	1.107	1.369	1.785
66	6.20	1.699	1.491	1.907	1.926
67	6.20	1.317	1.236	1.397	1.926
68	6.26	1.317	1.236	1.397	1.923
69	6.04	1.280	1.195	1.365	1.550
70	6.20	1.317	1.236	1.397	1.926

(*continued, next page*)

Table 4. (*continued*)

IUPAC	log K_{ow}	log E50	Lower CL	Upper CL	Empirical Bayes estimate
71	5.98	1.280	1.195	1.365	1.628
72	6.26	1.326	1.216	1.436	1.923
73	6.04	1.238	1.107	1.369	1.550
74	6.20	1.699	1.491	1.907	1.926
75	6.05	1.781	1.612	1.949	1.537
76	6.13	1.317	1.236	1.397	1.930
77	6.36	2.125	1.666	2.584	2.135
78	6.35	1.678	1.422	1.934	2.139
79	6.42	1.678	1.422	1.934	2.192
80	6.48	1.903	1.610	2.196	1.861
81	6.36	2.125	1.666	2.584	2.135
82	6.20	1.313	1.233	1.392	1.342
83	6.26	1.315	1.203	1.428	1.473
84	6.04	1.273	1.168	1.378	1.290
85	6.30	1.705	1.565	1.846	1.560
86	6.23	1.313	1.233	1.392	1.407
87	6.29	1.313	1.233	1.392	1.538
88	6.07	1.336	1.258	1.414	1.276
89	6.07	1.336	1.258	1.414	1.276
90	6.36	1.313	1.233	1.392	1.536
91	6.13	1.336	1.258	1.414	1.248
92	6.35	1.315	1.203	1.428	1.540
93	6.04	1.273	1.168	1.378	1.290
94	6.13	1.273	1.168	1.378	1.248
95	6.13	1.273	1.168	1.378	1.248
96	5.71	1.406	1.190	1.622	1.643
97	6.29	1.313	1.233	1.392	1.538
98	6.13	1.336	1.258	1.414	1.248
99	6.39	1.705	1.565	1.846	1.892
100	6.23	1.977	1.760	2.195	1.969
101	6.38	1.313	1.233	1.392	1.527
102	6.16	1.336	1.258	1.414	1.464
103	6.22	1.336	1.258	1.414	1.897
104	5.81	1.679	1.431	1.927	1.649
105	6.65	1.928	1.696	2.160	1.903
106	6.64	1.527	1.337	1.717	1.904
107	6.71	1.527	1.337	1.717	2.019
108	6.71	1.527	1.337	1.717	2.019
109	6.48	1.313	1.233	1.392	1.257
110	6.48	1.313	1.233	1.392	1.257
111	6.76	1.695	1.386	2.004	2.104
112	6.45	1.315	1.203	1.428	1.468
113	6.54	1.315	1.203	1.428	1.472
114	6.65	1.928	1.696	2.160	1.903
115	6.49	1.705	1.565	1.846	1.292
116	6.33	1.313	1.233	1.392	1.548
117	6.46	1.313	1.233	1.392	1.398
118	6.74	1.928	1.696	2.160	2.076
119	6.58	1.705	1.565	1.846	1.616
120	6.79	1.527	1.337	1.717	2.144
121	6.64	1.313	1.233	1.392	1.831
122	6.64	1.527	1.337	1.717	1.904
123	6.74	1.928	1.696	2.160	2.076
124	6.73	1.527	1.337	1.717	2.057
125	6.51	1.313	1.233	1.392	1.364
126	6.89	2.709	2.210	3.208	2.637
127	6.95	2.464	2.184	2.745	2.444
128	6.74	2.149	1.924	2.374	2.190
129	6.73	1.665	1.380	1.950	2.154
130	6.80	1.665	1.380	1.950	2.186
131	6.58	1.483	1.310	1.656	1.969
132	6.58	1.483	1.310	1.656	1.969
133	6.86	1.849	1.465	2.234	1.962
134	6.55	1.477	1.287	1.667	1.969
135	6.64	1.477	1.287	1.667	1.969
136	6.22	1.504	1.280	1.729	1.590
137	6.83	2.149	1.924	2.374	2.184
138	6.83	2.149	1.924	2.374	2.184
139	6.67	2.101	1.832	2.370	1.969
140	6.67	2.101	1.832	2.370	1.969
141	6.82	1.665	1.380	1.950	2.185
142	6.51	1.483	1.310	1.656	1.969
143	6.60	1.483	1.310	1.656	1.969
144	6.67	1.483	1.310	1.656	1.969
145	6.25	1.666	1.410	1.921	1.621
146	6.89	1.665	1.380	1.950	2.079
147	6.64	1.483	1.310	1.656	1.969
148	6.73	1.483	1.310	1.656	1.969
149	6.67	1.483	1.310	1.656	1.969
150	6.32	1.666	1.410	1.921	1.694
151	6.64	1.477	1.287	1.667	1.969
152	6.22	1.504	1.280	1.729	1.590
153	6.92	2.149	1.924	2.374	2.196
154	6.76	2.101	1.832	2.370	1.969
155	6.41	2.603	2.025	3.182	1.788
156	7.18	2.722	2.263	3.182	2.673
157	7.18	2.722	2.263	3.182	2.673
158	7.02	2.149	1.924	2.374	2.378
159	7.24	2.460	2.091	2.830	2.673
160	6.93	1.665	1.380	1.950	2.214
161	7.08	1.665	1.380	1.950	2.487
162	7.24	2.460	2.091	2.830	2.673
163	6.99	1.665	1.380	1.950	2.323
164	7.02	1.665	1.380	1.950	2.378
165	7.05	1.849	1.465	2.234	2.433
166	6.93	2.149	1.924	2.374	2.214
167	7.27	2.722	2.263	3.182	2.673
168	7.11	2.149	1.924	2.374	2.642
169	7.42	3.197	1.399	4.996	3.296
170	7.27	2.995	1.947	4.042	2.834
171	7.11	2.752	2.130	3.373	1.969
172	7.33	2.811	2.136	3.487	2.943
173	7.02	2.247	1.753	2.741	1.969
174	7.11	2.247	1.753	2.741	1.969
175	7.17	2.247	1.753	2.741	1.969
176	6.76	2.173	1.702	2.644	2.122
177	7.08	2.247	1.753	2.741	1.969
178	7.14	2.447	1.986	2.909	1.969
179	6.73	2.141	1.863	2.419	2.122
180	7.36	2.995	1.947	4.042	2.998
181	7.11	2.752	2.130	3.373	1.969
182	7.20	2.752	2.130	3.373	1.969
183	7.20	2.752	2.130	3.373	1.969
184	6.85	2.862	2.027	3.697	2.122
185	7.11	2.247	1.753	2.741	1.969
186	6.69	2.173	1.702	2.644	2.080
187	7.17	2.247	1.753	2.741	1.969
188	6.82	2.173	1.702	2.644	2.122
189	7.71	3.224	1.278	5.171	2.673
190	7.46	2.995	1.947	4.042	2.998
191	7.55	2.995	1.947	4.042	2.998
192	7.52	2.811	2.136	3.487	2.998
193	7.52	2.811	2.136	3.487	2.998
194	7.80	3.260	1.089	5.431	2.998
195	7.56	3.214	1.318	5.111	1.969
196	7.65	3.214	1.318	5.111	1.969
197	7.30	3.188	1.417	4.959	2.122
198	7.62	3.162	1.525	4.800	1.969
199	7.20	3.037	1.856	4.218	2.122
200	7.27	3.037	1.856	4.218	2.122
201	7.62	3.162	1.525	4.800	1.969
202	7.24	3.100	1.681	4.520	2.122
203	7.65	3.214	1.318	5.111	1.969
204	7.30	3.188	1.417	4.959	2.122
205	8.00	3.260	1.089	5.431	2.998
206	8.09	3.272	1.001	5.544	1.969
207	7.74	3.269	1.027	5.511	2.122
208	7.71	3.264	1.064	5.464	2.122
209	8.18	3.274	0.985	5.563	2.122

Abbreviations: IUPAC, International Union of Pure and Applied Chemistry; E50, the concentration producing an activity 50% above the control activity; CL, 95% confidence limit. The Bayes estimates of log E50 are also shown.

104.3 µM, respectively. Based on dose addition, the second order logistic model predicted 25.7, 32.0, and 69.3 µM for the E50 of these mixtures, respectively. Based on dose addition, the EBE predicted 35.4, 43.6, and 96.3 µM, respectively. The EBEs provide better predictions of mixture E50s; this suggests that the model is attributing to error important details of the structural relationship contained in the estimated log E50s.

Discussion

Adequacy of the concentration–activity model. A logistic model was fit to the activity–concentration data; the estimation of the maximum was unstable so an exponential model was used instead. This model is similar to the logistic, but it has no maximum. In fitting the exponential, we deleted data at high concentrations when the activity decreased with increasing concentration. This reversal in slope might indicate that the concentration–activity curve has reached a maximum and, for some congeners, the activity might never attain 50% above control. However, the congener could metabolize to a more soluble metabolite, justifying this deletion of data. On the other hand, congeners whose activity does not attain 50% above control might be considered inert with respect to the E50, but, for example, active with respect to the E10. Whether the congener is inert or weakly active is difficult to discern; however, this distinction is of less concern in mixture prediction.

Adequacy and validation of the model relating activity to structure. One objective of this study was to determine how reliable the predictions were for the untested congeners. The internal validation was very good with a Q^2 = 0.641. The use of a logistic model to describe the relationship between activity and chemical structure was very successful. The graph of the logistic model is a sigmoidal-shaped curve, so the range of predictions from this model is between two fixed parameters. Therefore, the prediction of extreme E50 values is avoided. The use of the logistic model causes the smallest E50 prediction to be near the average of other small E50s, rather than possibly being a value lower than any calculated E50. Because our model is empirical, this reluctance to extrapolate is reasonable. Logistic regression in this instance is very similar to discriminant analysis with two classes—active and inactive congeners.

There were indications of lack of fit of the model. In this case, one might predict the mixtures with the estimated E50s rather than the predictions from the logistic model. This strategy only works for tested congeners, and we still need a way to predict the E50s of the untested congeners. The EBE offers a compromise in this case. The EBE shrinks the estimated E50s toward the model-predicted E50s based on the amount of lack of fit. Then the EBEs for the untested congeners were obtained by straight line connections of these points when plotted versus log K_{ow}.

There are various ways this process could be improved. We want to evaluate both prediction methods using various types of PCB environmental mixtures before we decide which prediction method is best.

Comparison to other approaches for determining structure–activity relationships. It is enlightening to compare this approach with structure–activity modeling using Hansch analysis as described by Martin (*21*). Martin believes "the Hansch method is most suitable for a data set that has the following characteristics: 1) the compounds should be structural analogs that are identical in the structure of the pharmacopone, 2) all analogs should produce their biological effect by interacting with the same biological receptor(s), 3) it should be possible to derive quantitative measures of the physical properties of the analogs, 4) there should be enough compounds in the set that one can statistically examine a number of properties, . . . 5) the variation in potency between different analogs should be substantially larger than the error in measuring potency, and 6) the relevant physicochemical properties should be varied properly within the series."

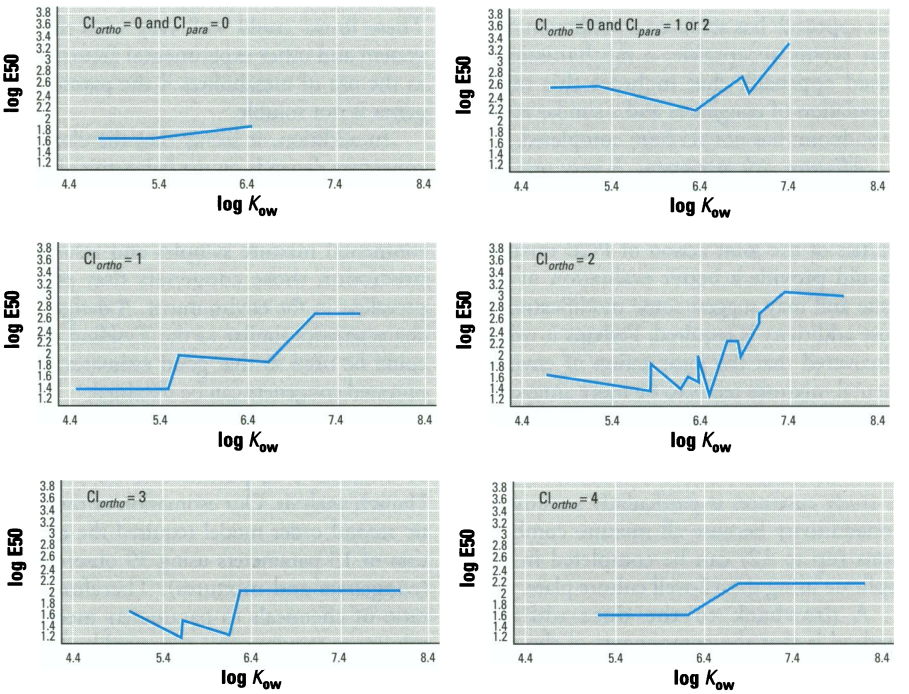

Figure 5. Bayes estimates of log E50 versus log K_{ow}. E50, the concentration producing an activity 50% above the control activity.

It is not clear that the 209 PCB congeners satisfy the Hansch assumptions. They do not all interact with the same receptor. This generally requires the dose–response slopes to be equal. We have fit the activity curves with the same slope, and this is a good approximation at this stage. More congeners should be tested to determine whether the slopes vary with subsets of the congeners. Thus, there is a lack of evidence to conclude that the slopes differ or that the shapes of the concentration–activity curves are not similar. On the other hand, there is increasing evidence that not all congeners interact with the same receptors. Therefore, PCB congeners may not satisfy this assumption, and Hansch analysis may not be appropriate. However, Hansch analysis may be appropriately applied to subsets of the 209 congeners once those subsets that satisfy the assumptions have been identified.

We have previously stated why we did not use mechanistically based variables in our model. However, Verhaar et al. (*22*) have shown an advantage of using mechanistically based variables in risk assessments of those complex mixtures in the environment whose chemical components are difficult to measure. Those interested in such approaches may find our predicted E50s useful to relate to mechanistically based variables.

Because this is an empirical approach, there are limitations. One cannot be as comfortable in the predictability from an empirical approach as from an approach that uses

mechanistically based variables. However, this approach has achieved a prediction error that seems acceptable; a mechanistically based approach that includes this target population of congeners and a similar activity endpoint has not yet been completed.

Mixture risk assessment. A very practical approach to predict PCB mixture activity is to use the dose-addition assumption. Future studies of the joint action of PCB mixtures should provide a better evaluation of this issue and may suggest some modification to dose additivity is required. Perhaps it will be discovered that the joint action of environmental mixtures tends to be more antagonistic on average. A simple correction could be introduced to account for this deviation from additivity. The use of this test system to study mixtures is particularly sensitive to nonadditivity because all mixture combinations (up to 12) can be completed in one replicate using the 12-well culture plate.

A drawback to the dose-addition method is that the components of the mixture need to be known, and this information can be difficult to obtain for some complex mixtures in the environment. However, if one knows the percent of chlorine in a PCB mixture, this and other knowledge could be used to calculate a probability interval for the E50 of an environmental PCB mixture that would be useful in determining risk. This calculation would be based on predicted E50s as we have provided and might require the assumption of a probability distribution for PCB congeners.

Congener selection. Tysklind et al. (*23*) discussed a procedure for selecting PCB congeners for use in quantitative structure–activity modeling. They restricted the congeners to the 154 tetra- through hepta-chlorinated congeners. Some restriction is necessary in view of the solubility problems encountered with most of the test systems. The congeners were described using 47 physicochemical variables, and principal component analysis reduced the number of variables to four orthogonal summary variables. The full 2^4 factorial design plus four center points required 20 congeners to be tested. These 20 congeners were divided into two groups. One group of 10 congeners was labeled the training set to be used for model development. The second group was labeled the validation set. The division occurred so that both sets of congeners typified the whole chemical domain of the four summary variables.

Given our goal of mixture prediction and the problems encountered with solubility, the approach considered by Tysklind et al. (*23*) does not seem appropriate. It is efficient to restrict testing to those congeners that are soluble, as their approach does. However, some sampling needs to be done to determine the limits of that region. These limits could well vary with the test system and other experimental factors, such as the use of metabolites if the parent congener is not soluble.

In addition, it seems that the choice of congeners depends on the balancing of several objectives. For example, we can test those congeners found in high percentages in environmental mixture samples and reserve for prediction those congeners that are less prevalent in the environment. A design with characteristics similar to the Tysklind design could be used so that good estimates of the untested congeners can be obtained.

Do more congeners need to be tested? Martin (*21*) has suggested that a 3 to 1 ratio of chemicals to parameters be used for Hansch analysis. Our estimation of the logistic second order model required the estimation of 13 parameters using 39 observations (congeners and mixtures). Therefore, our ratio of chemicals to parameters meets the minimum requirements suggested by Martin (*21*). In the area of general medical research, which is probably subject to more sources of prediction error than QSAR, Neter et al. (*24*) have suggested the use of a 6 to 1 ratio as a rule of thumb. Our more modest goal of prediction and description rather than variable selection may not require as large a ratio. Also, our use of the logistic model reduces the range of prediction. This reduction in range should limit the prediction error, which should translate into a less stringent requirement for precise estimates of the regression coefficients. The precision of these estimates is the reason for requiring a high ratio of chemicals to variables.

We think we have tested enough congeners. We have tried to externally validate various models based upon data from 17, 24, and 28 congeners. As the number of congeners fit to the model increased, the relative number of prediction errors decreased. The width of the confidence intervals as shown in Table 4 depend upon the fit of the model and how well the tested congeners span the space of the untested target congeners. Testing more congeners to better span the space of the target congeners is not likely to yield smaller confidence intervals due to insolubility limitations. Testing more congeners to justify the use of a more biologically based model such as a main effect model does not seem productive. Therefore, no further single congener testing using this test system is planned.

Further testing of mixtures is planned, however. We have tested a few other mixtures and expect to learn some new information about mixtures. When the model is fit to E50s arising mainly from mixture testing rather than from single congener testing, we initially expect a worsening of the fit of the model. This is due to information that the present model does not consider. For example, solubility might not have a dose-addition joint action and would interfere less in data consisting mainly of mixtures. After this phase has passed and enough mixtures have been tested, we expect the model to provide satisfactory predictions of mixtures.

Risk assessment implications. Although further research is needed to better understand whether this test system (PKC activation/translocation) has any role in the neurotoxicity of PCBs, data from this test system eventually may be used in terms of assessing the biological activity of PCB mixtures in a risk assessment. It is believed that this analysis of the data could provide benchmark estimates for mixtures following the method used by Crump (*25*) with some modifications. These modifications may include the method used to calculate confidence limits, the need for a lower bound on the slope estimate, and the use of a concentration causing a percentage of subjects to be affected rather than the E50.

Conclusion

This approach toward prediction and relating structure to activity is novel in several ways. The laboratory effort is reduced because our exponential model does not have the data requirement of estimating a maximum. The model has been validated, but there were indications of lack of fit. Also, the predictions of E50s for the three commercial mixtures based on dose addition were lower than the estimated E50s. Future mixture studies should indicate how the dose-additivity equation needs to be modified to predict various environmental mixtures. Our empirical approach is not limited to the assumptions of Hansch analysis. This approach, suitably modified, should be useful when applied to other test systems and should improve the ability to evaluate PCB congeners for hazard identification.

References

1. Erickson MD. Analytical Chemistry of PCBs. Boston, MA:Buttersworth, 1986.
2. Safe S, Safe L, Mullin M. Polychlorinated biphenyls: environmental occurrence and analysis. In: Polychlorinated Biphenyls (PCBs): Mammalian and Environmental Toxicology (Safe S, Hutzinger O, eds). Berlin:Springer-Verlag, 1987;1–13.
3. Safe S. Polychlorinated biphenyls (PCBs), dibenzo-*p*-dioxins (PCDDs), dibenzofurans (PCDFs), and related compounds: environmental and mechanistic considerations which support the development of toxic equivalency factors (TEFs). CRC Crit Rev Toxicol 21:51–88 (1990).
4. Seegal RF, Bush B, Shain W. Lightly chlorinated

ortho-substituted PCB congeners decrease dopamine in nonhuman primate brain and in tissue culture. Toxicol Appl Pharmacol 106:136–144 (1990).
5. Shain W, Bush B, Seegal RF. Neurotoxicity of polychlorinated biphenyls: structure–activity relationship of individual congeners. Toxicol Appl Pharmacol 111:33–42 (1991).
6. Kodavanti PRS, Ward TR, McKinney JD, Tilson HA. Increased [^{3}H]-phorbol ester binding in rat cerebellar granule cells by polychlorinated biphenyl mixtures and congeners: structure–activity relationships. Toxicol Appl Pharmacol 130:140–148 (1995).
7. Kodavanti PRS, Derr-Yellin EC, Mundy WR, Shafer TJ, Farmer JD, MacPhail RC, Tilson HA. Repeated exposure of adult rats to Aroclor 1254 alters motor activity and causes brain region specific changes in intracellular Ca^{2+} buffering and protein kinase C activity [abstract]. In: 26th Annual Meeting of the Society for Neuroscience, 16–21 November 1996, Washington, DC. Washington, DC: Society for Neuroscience, 1996;1910.
8. Davis JM, Svendsgaard DJ. U-Shaped dose-response curves: their occurrence and implications for risk assessment. J Toxicol Environ Health 30:71–83 (1990).
9. Kubinyi H. Quantitative structure–activity relationships IV. Non-linear dependence of biological activity on hydrophobic character: a new model. Arzneim-Forsh (Drug Res) 26(11):1991–1997 (1976).
10. Sabljiĉ A. Chemical topology and ecotoxicology. Sci Total Environ 109/110:197–220 (1991).
11. Kodavanti PRS, Ward TR, McKinney JD, Waller CL, Tilson HA. Increased [^{3}H]-phorbol ester binding in rat cerebellar granule cells and inhibition of $^{45}Ca^{2+}$ sequestration in rat cerebellum by polychlorinated diphenyl ether congeners and analogs: structure–activity relationships. Toxicol Appl Pharmacol 138:251–261 (1996).
12. Trilivas I, Brown JH. Increases in intracellular Ca^{2+} regulate the binding of [^{3}H]phorbol 12,13-dibutyrate to intact 1321N1 astrocytoma cells. J Biol Chem 264:3102–3107 (1989).
13. Vaccarino FM, Liljequist S, Tallman JF. Modulation of protein kinase C translocation by excitatory and inhibitory amino acids in primary cultures of neurons. J Neurochem 57:391–396 (1991).
14. Mattson MP. Evidence for the involvement of protein kinase C in neurodegenerative changes in cultured human cortical neurons. Exp Neurol 112:95–103 (1991).
15. Eboli ML, Ciotti MT, Mercanti D, Calissano P. Differential involvement of protein kinase C in transmitter release and response to excitatory amino acids in cultured cerebellar neurons. Neurochem Res 18:133–138 (1993).
16. Hawker DW, Connell DW. Octanol–water partition coefficients of polychlorinated biphenyl congeners. Environ Sci Technol 22:382–387 (1988).
17. Frame GM, Cochran JW, Bøwadt SS. Complete PCB congener distributions for 17 Aroclor mixtures determined by 3 HRGC systems optimized for comprehensive, quantitative, congener-specific analysis. J High Resol Chromatogr 19:657–668 (1996).
18. Draper NR, Smith H. Applied Regression Analysis. New York:John Wiley & Sons, 1981.
19. Svendsgaard DJ, Hertzberg RC. Statistical methods for the toxicological evaluation of the additivity assumption as used in the Environmental Protection Agency chemical mixture risk assessment guidelines. In: Toxicology of Chemical Mixtures: Case Studies, Mechanisms, and Novel Approaches (Yang RSH, ed). San Diego, CA:Academic Press, 1994;599–641.
20. Bryk AS, Raudenbush SW. Hierarchical Linear Model: Application and Analysis Methods. Newbury Park, CA:Sage Publications, 1992.
21. Martin YC. Studies of relationships between structural properties and biological activity by Hansch analysis. In: Structure–Activity Correlation as a Predictive Tool in Toxicology: Fundamentals, Methods, and Applications (Golberg L, ed). Washington, DC:Hemisphere Publishing Corporation, 1983;77–92.
22. Verhaar HJM, Busser FLM, Hermans JLM. A surrogate parameter for the baseline toxicity content of contaminated water. Environ Sci Technol 29:726–734 (1995).
23. Tysklind M, Andersson P, Haglund P, Van Bavel B, Rappe C. Selection of polychlorinated biphenyls for use in quantitative structure–activity modelling. SAR and QSAR in Environmental Research 4:11–19 (1995).
24. Neter J, Wasserman W, Kutner MH. Applied Linear Statistical Models Regression, Analysis of Variance, and Experimental Designs. 3rd Ed. Burr Ridge, IL:Irwin, 1990.
25. Crump K. A new method of determining daily intakes. Fundam Appl Toxicol 4:854–871 (1984).